

Panelanalyse im Überblick

Faulbaum, Frank

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Faulbaum, F. (1988). Panelanalyse im Überblick. *ZUMA Nachrichten*, 12(23), 26-44. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-210016>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Panelanalyse im Überblick

Der nachfolgende Aufsatz gibt einen kurzen Überblick über den Stand der Modellierung von Panel-daten und zeigt einige wichtige Entwicklungstendenzen in diesem Bereich auf.

1. Einleitung

Die Anzahl der empirischen Studien, in denen Paneldaten erhoben werden (Panelstudien), ist in den letzten Jahren ohne Zweifel sprunghaft gestiegen. Beispiele für große europäische Panelstudien im sozio-ökonomischen Bereich sind das vom Sonderforschungsbereich 3 "Mikroanalytische Grundlagen der Gesellschaftspolitik" der Universitäten Frankfurt und Mannheim in Zusammenarbeit mit dem Deutschen Institut für Wirtschaftsforschung durchgeführte und von der Deutschen Forschungsgemeinschaft geförderte Sozio-ökonomische Panel (vgl. Hanefeld 1984; Hanefeld 1987; Krupp/Hanefeld 1987), das schwedische Projekt "Household Market and Nonmarket Activities" (vgl. Klevmarken 1984), das niederländische Sozio-ökonomische Panel (vgl. Van de Stadt 1987; Van de Stadt et al. 1986, 1987) sowie das luxemburgische "Panel Socioeconomique Lorraine et Luxembourg" des Centre d'Etudes de Populations, de Pauvete et de Politiques Socio-Economiques (CEPS) in Walferdange (Luxemburg) (vgl. z.B. Schaber et al. 1982). Vorbild für viele Aspekte dieser großen Studien war die U.S. Panel Study of Income Dynamics (PSID; vgl. Duncan/Morgan 1984). 1986 wurde unter Leitung von Greg J. Duncan am CEPS das Panel Data Comparability Project (PACK) begonnen, das die verschiedenen internationalen Panelprojekte koordinieren und vergleichende internationale Sozialforschung erleichtern soll.

Bekannte deutsche Panelerhebungen sind die Untersuchungen zur beruflichen Verbleibsforschung des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit in Nürnberg, die Hochschulabsolventenstudie des Wissenschaftlichen Zentrums für Berufs- und Hochschulforschung der Gesamthochschule Kassel, die Panelerhebungen des Hochschul-Informations-Systems in Hannover u.a. Im Bereich der Wahlforschung gehören Paneluntersuchungen inzwischen zum methodologischen Standardrepertoire.

Neben diesen Studien mit praxisorientierten gesellschafts-, bildungs- und/oder wirtschaftspolitischen Zielsetzungen, deren Aufgabe zu einem guten Teil auch darin besteht, Erkenntnisse über gesellschaftliche Veränderungen und Entwicklungen zu sammeln, die gegebenenfalls auch als Planungsunterlagen für Entscheidungen in bestimmten gesellschaftlichen Bereichen verwendet werden können, gibt es eine ganze Reihe weiterer Studien, die eher als quasi-experimentelle Studien einzustufen sind. Diese haben das Ziel, theoretische oder methodische Fragestellungen zu untersuchen. Hierzu gehören z.B. die Begleituntersuchung zur Volkszählung 1987 des Statistischen Bundesamtes oder

die ZUMA-Test-Retest-Studie zum ALLBUS 1984 (vgl. die Beiträge in Bohrnstedt/Mohler/Müller 1987).

Verbindet man mit dem Begriff "Paneldaten" wiederholte Beobachtungen des gleichen Merkmals an gleichen Individuen über mehrere Zeitpunkte, so ist klar, daß die meisten Panelstudien nur zu einem Teil Paneldaten erheben. Es gibt immer einen gewissen Anteil von Fragen zu Themenbereichen, die von Zeitpunkt zu Zeitpunkt variieren. In vielen Studien wird auch versucht, fehlende Ereignisdaten zu ergänzen, die sich auf den Zeitraum zwischen den Wellen beziehen. Dies geschieht durch retrospektive Fragen, die sich etwa auf den Zeitpunkt der eingetretenen Arbeitslosigkeit, den Zeitpunkt des Berufswechsels oder auf die Zeitdauer dazwischenliegender Lebensereignisse (life events) beziehen. Damit kommen für die Auswertung von Panelstudien nicht nur Methoden zur Analyse von Paneldaten, sondern auch solche zur Analyse von Ereignisdaten zur Anwendung (vgl. z.B. Arminger 1984; Blossfeld/Hamerle/Mayer 1986; Carrol 1982; Diekmann in diesem Heft; Diekmann/Mitter 1984; Heckman/Singer 1984, 1986; Tuma/Hannan 1984). Die Anzahl der Meßzeitpunkte (Wellen) variiert von Studie zu Studie, und es ist nicht auszuschließen, daß bei großen Studien, wie dem Sozio-ökonomischen Panel, schließlich nicht mehr Paneldaten, sondern Zeitreihendaten vorliegen, was die Anwendung einer weiteren Klasse statistischer Verfahren ermöglicht.

Mit der Erhebung von Paneldaten können sich ganz verschiedene theoretische und/oder praktische Ziele verbinden. Das primäre Ziel besteht in der Beschreibung, Erklärung und Prognose von Veränderungen. Darüber hinaus gibt es eine Reihe weiterer methodischer Ziele, die sich nur in Studien mit mehr als zwei Meßzeitpunkten realisieren lassen. Dazu gehören insbesondere die separate Bestimmung von Reliabilität und Stabilität von Messungen (vgl. Heise 1969; Wiley/Wiley 1970), die Schätzung item-spezifischer Einflüsse (vgl. Raffalovich/Bohrnstedt 1987), die statistische Kontrolle unbeobachteter individueller und/oder zeitlicher Heterogenität, wie sie in den statischen oder dynamischen Zufallskomponentenmodellen vorgenommen wird (vgl. z.B. Anderson/Hsiao 1982; Bhargava/Sargan 1983; Chamberlain 1982, 1984; Hsiao 1986) oder die Vermeidung von Fehlspezifikation durch ausgelassene (nicht erhobene) Variablen.

Unter den Auswertungsverfahren in der empirischen Panelforschung überwiegen Verfahren der statistischen Modellierung. Hierbei werden Parameter von Wahrscheinlichkeitsverteilungen beobachteter oder latenter Variablen, wie etwa Erwartungswerte oder Kovarianzen, als Funktionen von Strukturparametern und beobachtbaren oder latenten Residualvarianzen und -kovarianzen dargestellt. In letzter Zeit wird versucht, auch die höheren Momente von Verteilungen, wie die Schiefe, zu modellieren (vgl. Bentler 1983).

Die Anwendung mehr explorativer Verfahren, bei denen die Entdeckung und Identifikation von Strukturen und Substrukturen sowie die Hypothesengenerierung im Vordergrund stehen, findet man nur selten. So bildet etwa die Arbeit von Van der Heijden (1987), in der korrespondenzanalytische Verfahren zur Identifikation zeitlicher Abhängigkeitsstrukturen herangezogen werden, eindeutig die Ausnahme. Es bleibt abzuwarten, welche neuen heuristischen Möglichkeiten der Strukturidentifikation und Hypothesenerzeugung sich aus der Anwendung der dynamischen Graphik (vgl. Becker/Cleveland/Wilks 1987; Huber 1985; Sawitzky 1988; Young 1988) und den computerunterstützten logischen Verfahren der Entdeckung statistisch relevanter Hypothesen ergeben (vgl. Hajek/Havranek 1978a, 1978b). Auch die Anwendung voraussetzungsloserer Skalierungsverfahren für Longitudinaldaten, wie z.B. longitudinale Guttman-Simplizes (vgl. Collins/Cliff 1985), haben in der Panelanalyse bis jetzt noch keinen Einzug gehalten. Als Überblick über die gegenwärtig möglichen Modellierungstechniken im Bereich der Panelanalyse können die im Handel befindlichen Einführungen (vgl. z.B. Markus 1979; Kessler/Greenberg 1981; Plewis 1985) leider nicht uneingeschränkt empfohlen werden.

2. Formen der Modellierung von Paneldaten

Die Klassifikation von Modellen zur Analyse von Paneldaten kann nach unterschiedlichen Gesichtspunkten erfolgen, denn sie hängen in ihrer mathematischen Gestalt und den zur Verfügung stehenden statistischen Analyseverfahren von gewissen theoretischen Vorentscheidungen ab. Grundsätzlich lassen sich Paneldaten im Rahmen von statischen oder dynamischen Modellen, von univariaten oder multivariaten Modellen, von Modellen mit diskreter oder kontinuierlicher Zeitachse und von Modellen mit diskreten und/oder kontinuierlichen endogenen Variablen erklären. Im Falle dynamischer Modelle wird die Realisierung der Werte (Zustände) mindestens einer endogenen Variablen zum Zeitpunkt t von der Realisierung der Werte (Zustände) zu vorherigen Zeitpunkten, meistens jedoch nur zum Zeitpunkt $t-1$, abhängig gemacht (Einführung verzögerter bzw. "lagged" endogener Variablen).

Im allgemeinen Fall wird versucht, Hypothesen über die zeitliche Veränderung von k Variablen (z.B. k Einstellungsvariablen) y_k ($k=1, \dots, K$) zu Zeitpunkten t ($t=1, \dots, T$) und ihre funktionalen Abhängigkeiten von exogenen Einflußstrukturen auf formale Modelle abzubilden. Die verwendeten Modelle sind zu meist mathematischer (also z.B. nicht logischer), und hier vor allem statistischer Natur. Empirisch vorliegen haben wir an N Einheiten (z.B. Personen, Querschnitte etc.) zu den Zeitpunkten t bezüglich der Variablen y_k vorgenommene Beobachtungen (Messungen) y_{itk} ($i=1, \dots, N$; $t=1, \dots, T$; $k=1, \dots, K$) sowie gegebenenfalls Beobachtungen gewisser exogener Variablen. Über die bedingten oder unbedingten unbekannten Verteilungen der y_{itk} werden bestimmte statistische Hintergrundannahmen gemacht, die mehr oder weniger plausibel sein können und damit eine erste Quelle für Fehlspezifikation darstellen. Manch-

mal werden auch Terme für feste oder zufällige Effekte unbeachteter individueller oder zeitlicher Heterogenität eingeführt.

Häufig werden die endogenen und/oder exogenen Variablen auch als latent angenommen. Obgleich sich die Anwendungen von Modellen mit latenten Variablen in der Regel auf Modelle mit diskreter Zeitachse beziehen, gibt es auch Beispiele für die Modellierung mit angenommener kontinuierlicher Zeitachse (vgl. Arminger 1986; Möbus 1986), wobei die latenten Variablen allerdings als kontinuierlich vorausgesetzt werden. Ein latentes Analogon zu den von Coleman (1964, 1968, 1981) betrachteten Modellen diskreter Veränderungen in kontinuierlicher Zeit ist dem Autor nicht bekannt.

Aus Gründen begrifflicher Klarheit empfiehlt es sich, zwischen latenten Antwortvariablen und latenten Konstruktvariablen zu unterscheiden. Alle Modelle für limitierte abhängige Variablen (limited dependent variables; vgl. Amemiya 1981, 1984, 1985; Dhrymes 1984; Maddala 1983; McFadden 1984) basieren auf der Annahme von in bestimmter Art und Weise verteilten latenten Antwortvariablen y^* , wobei die Beobachtungen y durch einen stochastischen Auswahlmechanismus erzeugt werden. Alle Modelle mit diskreten ordinalen Variablen, wie Wahrscheinlichkeits-, Logit-, Probit-, Tobit -, Loglineare Modelle einschließlich von Modellen für zensierte Variablen, lassen sich über die Annahme latenter Antwortvariablen reformulieren und als Spezialfälle von Modellen für limitierte abhängige Variablen auffassen. Latente Konstruktvariablen hingegen entsprechen nur über beobachtbare Indikatoren indirekt meßbaren theoretischen Konstrukten. Während alle "klassischen" Ansätze der Analyse von Kausalmodellen mit latenten Variablen vor allem an latente Konstruktvariablen dachten (vgl. z.B. Jöreskog 1969, 1978; Bentler 1980; Bentler/Weeks 1979) bezieht Muthén (1982, 1983, 1984) diese Unterscheidung in seinen Modellansatz explizit mit ein.

Die meisten Modelle zur Analyse von Paneldaten können unter zwei verschiedenen Aspekten interpretiert werden. Zum einen gibt es Ansätze, die beim Entwurf von Modellen primär an die Beschreibung von Variablenbeziehungen denken, zum anderen solche, die primär die Erzeugung zeitabhängiger Daten durch stochastische Systeme modellieren wollen. Selbst wenn wir uns bewußt sind, daß diese Unterscheidung bis zu einem gewissen Grad artifiziell ist, liegen doch zwei unterschiedliche Orientierungen vor. Im zweiten Fall muß besonderer Wert auf die Koordination der Beobachtungsdynamik mit der Prozeßdynamik gelegt werden. Im ersten Fall dagegen stehen Fragen der Stabilität und Veränderung im Vordergrund, ohne daß der Versuch gemacht wird, diese Veränderungen als Konsequenzen eines Prozeßmodells darzustellen. Unter rein mathematisch-statistischen Gesichtspunkten kommen sehr häufig für beide Orientierungen ähnliche Modelle in Frage.

Übersicht 1: Modelle für Paneldaten

QUANTITATIVE ABHÄNGIGE VARIABLE

DISKRETE ABHÄNGIGE VARIABLE

BEOBACHTETE ENDOGENE VARIABLE

- DISKRETE ZEIT**
1. Statische und dynamische lineare Modelle mit festen oder zufälligen Effekten (random component models bzw. random coefficient models) zur Erfassung unbeobachteter individueller und/oder zeitlicher Heterogenität (vgl. Anderson/Hsiao 1982; Arminger/Müller in Vorbereitung; Bhargava/Sargan 1983; Chamberlain 1982, 1984; Hsiao 1986; Kessler/Greenberg 1981)
 2. Strukturgleichungsmodelle (mehrere abhängige Variablen, interdependente Systeme; vgl. Duncan 1975)
 3. Stochastische Prozesse in diskreter Zeit (abzählbarer Indexmenge) mit kontinuierlichen Zustandsräumen wie Autoregressive Prozesse, Markoff-Simplizes (vgl. Anderson 1979; Anderson/Goodman 1957; Frederiksen/Rotondo 1979. Zur Darstellung durch Strukturgleichungsmodelle vgl. Jöreskog 1970, 1974, 1979, 1981)

BEOBACHTETE ENDOGENE VARIABLE

1. Modelle mit limitierten abhängigen Variablen (z.B. Wahrscheinlichkeits-, Logit-, Probit-, Tobit-Modelle, Log-lineare Modelle, Modelle mit zensierten abhängigen Variablen; vgl. Amemiya 1981; Dhrymes 1984; Maddala 1983; McFadden 1984)

2. Dynamische Wahlmodelle (Heckman 1978a, 1978b, 1981)

3. Stochastische Prozessmodelle mit diskreten Zustandsräumen (Beobachtete Markoff-Ketten, Mover-Stayer-Modelle; vgl. Anderson 1979; Blumen/Kogan/McCarthy 1955, 1966; Bishop/Fienberg/Holland 1975; Kap. 7; Singer 1982; Singer/Spielman 1976a)

KONTINUIERLICHE

ZEIT

1. Stochastische Differentialgleichungen (vgl. Arminger 1986; Hamerle/Magj/Singer 1988; Möbus/Magj 1983; Möbus 1986; Soong 1973; Tuma/Hannan 1984)

Coleman 1964, 1968, 1981; Chung 1960; Singer/Spielman 1976b)

Fortsetzung Übersicht 1

| | | DISKRETE ABHÄNGIGE VARIABLE |
|----------------------|---|---|
| DISKRETE ZEIT | QUANTITATIVE ABHÄNGIGE VARIABLE | |
| | LATENTE ENDOGENE VARIABLE | LATENTE DISKRETE VARIABLE |
| | <p>1. Latente Antwortvariablen: Siehe Modelle mit diskreten abhängigen Variablen</p> <p>2. Latente Konstruktvariablen: Strukturgleichungsmodelle mit latenten Konstruktvariablen</p> <p>A. Modelle mit kontinuierlichen Indikatoren (vgl. Aigner/Hsiao/Kapteyn/Vansbeek 1984; Bentler 1980, 1983; Bentler/Weeks 1980; Jöreskog 1973, 1978; McArdle/McDonald 1984; McDonald 1980)</p> <p>B. Modelle mit diskreten ordinalen Indikatoren (Muthén 1978, 1982, 1983, 1984; Muthén/Kaplan 1985; Olsson 1979; Olsson/Drasgow/Dorans 1982)</p> <p>C. Latente autoregressive Prozesse, Quasi-Markoff-Simplizes Jöreskog 1974, 1979)</p> | <p>1. "Latent Class"-Modelle (vgl. Dayton/Macready 1983; Langeheine 1987; McCutcheon 1987), Latente Markoffketten (Hagenaars 1987; Van de Pol/de Leeuw 1986), Gemischte diskrete Prozeßmodelle (Van de Pol/Langeheine 1988)</p> |
| KONTINUIERLICHE ZEIT | | |
| | <p>1. Stochastische Systeme mit kontinuierlicher latenter Dynamik (Arminger 1986; Möbus 1986)</p> | |

Die Übersicht 1 vermittelt einen Überblick über die Modellansätze der Panelanalyse. Dabei sind einige Typen aus Gründen der Akzentuierung nebengeordnet, obwohl sie unter rein mathematisch-statistischen Gründen eher über- bzw. untergeordnet sind. Die Literaturangaben beziehen sich einerseits auf allgemein einführende Literatur und andererseits auf spezielle Anwendungen in der Panelanalyse.

Viele Panelmodelle sind bisher in den empirischen Sozialwissenschaften kaum zur Anwendung gekommen. Dies gilt insbesondere für die in der Publikation von Hsiao (1986) extensiv beschriebenen Modelle mit festen und/oder zufälligen Effekten, die bisher vorwiegend in der Ökonomie eine Rolle gespielt haben (vgl. z.B. Bhargava/Sargan 1983). Ein Modell mit zufälligen Effekten hat in seiner dynamischen Form mit zeitinvarianten und zeitvariierenden Variablen sowie einem zufälligen Effekt für unbeobachtete individuelle Heterogenität die folgende Form:

$$y_{it} = \gamma y_{it-1} + \rho' z_i + \beta' x_{it} + v_{it} \quad (i=1, \dots, N; t=1, \dots, T)$$

mit: $|\gamma| < 1$, $v_{it} = \alpha_i + u_{it}$, wobei:

- α_i : individuenspezifischer zufälliger Effekt;
- u_{it} : Zufallsfehler;
- z_i : $p \times 1$ -Vektor zeitinvarianter exogener Variablen;
- ρ' : $1 \times p$ -Vektor der Einflußgrößen für die z_i ;
- x_{it} : $q \times 1$ -Vektor zeitvariierender exogener Variablen;
- β' : $1 \times q$ -Vektor der Einflußgrößen für die x_{it} .

Es wird angenommen, daß die Fehler untereinander und mit den exogenen Variablen unkorreliert sind.

Die zögerliche Rezeption dieser Modelle, wie auch einer Reihe anderer von seiten der empirischen Sozialwissenschaft, hat wohl nicht nur mit der Tatsache zu tun, daß es sich hier um rein quantitative Modelle handelt. Vielmehr scheint auch eine Rolle zu spielen, daß die notwendigen Schätzalgorithmen in den großen Standard-Statistik-Paketen nicht immer angeboten werden, so daß man nicht selten darauf angewiesen ist, die Schätzung selber zu programmieren. Mittlerweile sind aber Programme auf dem Markt (z. B. GAUSS, vgl. Edlefsen/Jones 1986), die das Programmieren von Schätzern sehr erleichtern. Manche dieser Modelle lassen sich aber auch mit Programmen zur Analyse von Strukturgleichungsmodellen darstellen und überprüfen (vgl. die LISREL-Beispiele in Arminger/Müller in Vorbereitung).

Die Überprüfung der Annahmen, die diesen Modellen zugrundeliegen, wie Unkorreliertheit von Fehlern und exogenen Variablen, wird durch die Weiterentwicklung von Spezifikationstests – vor allem vom Hausman-Typ – ermöglicht

(vgl. Arminger 1987; Arminger/Schönberg 1987; Hausman 1978, 1983; Hausman/Taylor 1981; White 1985). Ein Spezifikationstest dieses Typs basiert auf der Idee, daß unter der Nullhypothese (keine Fehlspezifikation) ein konsistenter, asymptotisch normaler und effizienter Schätzer existiert, der unter der Alternativhypothese (Fehlspezifikation) nicht erwartungstreu und konsistent ist. Es wird dann ein alternativer Schätzer gesucht, der unter der Nullhypothese nicht asymptotisch effizient ist, der aber durch die Fehlspezifikation in seinen Eigenschaften nicht beeinflußt wird. Geprüft wird, ob die Differenz dieser beiden Schätzer in der Wahrscheinlichkeit gegen 0 geht. Bei korrekter Spezifikation dürfen sie sich nicht unterscheiden.

3. Für die Modellierung von Paneldaten wichtige methodische Entwicklungen

Für die Modellierung von Paneldaten sind einige methodische Fortschritte relevant, die nicht nur Paneldaten im besonderen, sondern die Spezifikation von empirischen Modellen allgemein betreffen. Diese Entwicklungen sind insbesondere für Paneldaten von Bedeutung, die im Rahmen der Umfrage- und Einstellungsforschung erhoben wurden.

Neben den im letzten Abschnitt erwähnten Spezifikationstests sind es vor allem Fortschritte in der Behandlung fehlender Werte und Entwicklungen, die man unter der Überschrift "voraussetzungsärmere Ansätze" subsumieren könnte.

3.1 Behandlung fehlender Werte

Einen guten Überblick über die Behandlung fehlender Werte geben Little/Rubin (1987). Wie man mit fehlenden Werten umgeht, richtet sich, wie bei Querschnittsuntersuchungen danach, ob der Mechanismus, der die fehlenden Werte erzeugt hat, (missing data mechanism) ignorierbar (ignorable) ist oder nicht. Dieser Mechanismus ist nur dann ignorierbar, wenn die Daten im Rahmen einer erfolgreich durchgeführten Wahrscheinlichkeitsauswahl realisiert wurden. In diesem Fall ist der Mechanismus zur Erzeugung fehlender Werte unter der Kontrolle des Statistikers. Fehlende Werte sind in diesem Fall alle Werte, die im Rahmen des Stichprobendesigns nicht in die Stichprobe kamen.

In allen anderen Fällen müssen wir gewisse Annahmen über den missing data-Mechanismus treffen, wodurch dieser nicht mehr ignoriert werden kann. Jede Behandlung fehlender Werte im Rahmen von statistischen Analyseverfahren macht implizite Annahmen über diesen Mechanismus.

Die Behandlung fehlender Werte ist kein spezielles Problem der Panelanalyse, und so sind alle zur Diskussion stehenden Verfahren auch bei der Analyse von Paneldaten anwendbar. Diese Verfahren können sein: Verfahren, die auf vollständig erfaßten Einheiten beruhen (Herausnehmen aller Einheiten mit mindestens einem fehlenden Wert), Imputationsverfahren, bei denen die fehlenden Werte nach gewissen Standardverfahren, wie z.B. Regressionsverfahren, er-

setzt werden, Gewichtungsverfahren, bei denen eine Gewichtung vorgenommen wird, die den Effekt fehlender Werte korrigieren soll, und schließlich Verfahren, die auf expliziter Modellierung der fehlenden Werte beruhen.

Entscheidend für die Wahl des Verfahrens ist, ob die fehlenden Werte eine nach Zufall gezogene Substichprobe der Stichprobe darstellen (missing at random). In dem Fall, wo wir eine vollständig beobachtete Variable und eine weitere, mit fehlenden Werten ausgestattete Variable haben, empfiehlt sich auch eine Einteilung danach, ob die Antwortwahrscheinlichkeit von X und Y unabhängig ist ("completely missing at random", CMAR), ob sie nur von X abhängen ("missing at random", MAR) oder ob sie sowohl von X als auch von Y abhängen. Bei CMAR stellen die beobachteten Werte von Y eine Zufalls-Substichprobe der gezogenen Stichprobe dar, bei MAR stellen die Y-Werte nur eine Zufallsstichprobe innerhalb der durch die X-Werte definierten Klassen dar.

Die gängigen Imputationsverfahren und auch die auf einem expliziten statistischen Modell und auf Likelihood-Verfahren beruhenden Ansätze der Parameterschätzung bei fehlenden Daten setzen entweder CMAR oder MAR voraus. Dabei stehen sowohl Methoden zur Schätzung von Mittelwerten und Varianzen als auch zur Schätzung von Regressionskoeffizienten zur Verfügung. Im Fall der bei Paneldaten häufig auftretenden monotonen Datenmuster (Ausfälle erfolgen zwischen den Wellen), vereinfacht sich die Schätzung der Mittelwerte und Kovarianzen, da man in diesem Fall die gemeinsame Verteilung der Variablen faktorisieren kann (vgl. Little/Rubin 1987: 107; Marini/Olsen/Rubin 1980; Rubin 1987: 170). Arminger/Müller (in Vorbereitung) geben ein Beispiel für die Programmierung des Schätzalgorithmus in GAUSS.

Die Annahme von CMAR bzw. MAR dürfte allerdings in den meisten Fällen eine Illusion darstellen, so daß in diesem Fall der Auswahlmechanismus nicht ignoriert, sondern explizit modelliert werden muß. Hierbei kann auf Modelle mit limitierten abhängigen Variablen zurückgegriffen werden (vgl. Hsiao 1986: 198; Little/Rubin 1987: 223).

Spezielle Situationen entstehen, wenn für die Panelausfälle nach Zufall entsprechend viele neue Fälle substituiert werden (rotierende Stichproben). In diesem Fall können die Verfahren zur Analyse vollständiger Paneldaten entsprechend erweitert werden (vgl. hierzu Hsiao 1986: 193).

Von besonderer Bedeutung ist die Charakterisierung von Panelausfällen. Liegen genügend Ausfälle zu spezifischen Zeitpunkten vor, so kann man versuchen, im Rahmen von Gruppenvergleichen zwischen ausgefallenen Einheiten und nicht ausgefallenen Einheiten zu bestimmten Zeitpunkten Verhaltensunterschiede im vorangehenden Zeitverlauf zu überprüfen, so z.B. Unterschiede in den Stabilitäten, Reliabilitäten etc. Faulbaum (1987) vergleicht eine Gruppe

von in der dritten Welle ausgefallenen Befragten mit der Gruppe der in der dritten Welle verbliebenen Befragten hinsichtlich der latenten Mittelwerte einer latenten Einstellungsvariablen in den vorherigen beiden Wellen.

3.2 Voraussetzungsarme Ansätze

Nimmt man als Referenzpunkt Modelle mit quantitativen abhängigen (beobachteten oder latenten) Variablen mit – im Falle latenter Variablen – kontinuierlichen Indikatoren, deren Parameter unter Normalverteilungsbedingungen geschätzt werden müssen, lassen sich wesentliche Fortschritte ausmachen, die die Abschwächung der einen oder anderen Bedingung betreffen. Diese Fortschritte betreffen vor allem Strukturgleichungsmodelle mit oder ohne latente Variablen, aber auch andere Modellierungsformen. Bei den erwähnten Abschwächungen handelt es sich vor allem um die Abschwächung des Meßniveaus der abhängigen Variablen, die Abschwächung von Verteilungsvoraussetzungen und die Verallgemeinerung von Parameterrestriktionen.

3.2.1 Abschwächung des Meßniveaus

Wesentliche Fortschritte sind bei Modellen mit kategorialen abhängigen Variablen erzielt worden, was nicht zuletzt damit zusammenhängt, daß man alle Ansätze einheitlich als Spezialfälle von Schwellenwertmodellen (Modelle mit abhängigen limitierten Variablen) auffassen kann (s. Abschnitt 2). Den komplexen Fall eines dynamischen Modells mit einer dichotomen abhängigen Variablen beschreibt Heckman (1978a, 1978b, 1981):

$$y_{it}^* = \beta'x_{it} + \sum_{j=1}^{t-1} \gamma_j y_{it-j} + \phi \sum_{s=1}^{t-1} \sum_{j=1}^s y_{it-j} + v_{it} \quad (i=1, \dots, N; t=1, \dots, T)$$

und

$$y_{it} = \begin{cases} 1, & \text{falls } y_{it}^* > 0, \\ 0, & \text{falls } y_{it}^* \leq 0. \end{cases}$$

Hier ist y^* eine latente Antwortvariable. Es wird angenommen, daß der Fehlerterm v_{it} von den exogenen Variablen unabhängig ist und über i unabhängig verteilt ist. γ_j mißt den Effekt, den ein 1 Perioden zurückliegendes Ereignis auf die gegenwärtigen Werte von y_{it}^* hat. ϕ mißt den kumulativen Effekt, den Ereignisse zu unterschiedlich lang zurückliegenden Ereignissen auf die Gegenwart haben.

Besondere Weiterentwicklungen mit Folgen für die Analyse von Paneldaten sind auch bei der Modellierung mit latenten Variablen mit kategorialen Indikatoren und diskreten latenten Variablen (latenten Klassen) festzustellen. Im ersten Fall ist in erster Linie auf den Ansatz von Muthén zu verweisen. Muthén (1982, 1983, 1984) schlägt in Erweiterung der Ansätze von Muthén (1978) sowie Muthén/Christoffersson (1981) ein allgemeines Strukturgleichungsmodell vor.

chungsmodell für kontinuierliche normalverteilte latente Variablen mit ordinalen kategorialen (Likert-) und kontinuierlichen Variablen als Indikatoren vor. Dieses Modell beruht im wesentlichen darauf, daß zwischen latente Konstrukte und beobachtete Indikatoren latente Antwortvariablen geschoben werden, die mit den beobachteten Variablen im Rahmen von Schwellenwertmodellen miteinander verknüpft werden. Der Ansatz von Muthén umfaßt also, neben den auch in anderen Ansätzen analysierbaren Modellen mit kontinuierlichen Indikatoren, zusätzlich Modelle mit limitierten abhängigen Variablen, wobei sich auch Probit- und Tobit (Tobin's Probit)-Modelle (zensierte Regressionsmodelle; vgl. Tobin 1958) als Meßmodelle spezifizieren lassen.

Der Ansatz von Muthén basiert auf der Analyse von Kovarianzstrukturen. Sein Programm LISCOMP (vgl. Muthén 1987) gestattet die Analyse von Modellen mit latenten Konstruktvariablen und Indikatoren der oben beschriebenen Art.

Im Bereich der Analyse latenter Klassen (LCA; einführende Darstellungen: Formann 1984; McCutcheon 1987) gibt es eine Reihe innovativer Anwendungen auf die Modellierung zeitabhängiger Daten (vgl. z.B. Clogg 1981; Dayton/Macready 1983; Bye/Schechter 1986; Van de Pol/DeLeeuw 1986; Van de Pol/Langeheine 1988; Hagenaars 1987, 1988; Hagenaars/Luijkx 1987, zit. nach Hagenaars 1988). Obgleich die dort diskutierten Modellstrukturen bereits in den klassischen Arbeiten von Lazarsfeld (vgl. Lazarsfeld 1950, 1960, 1965; Lazarsfeld/Henry 1968) und Wiggins (Wiggins 1951, 1973) beschrieben und formal dargestellt wurden, mußte erst die Entwicklung entsprechender Schätzalgorithmen abgewartet werden, um diese Modelle auch für die praktische Anwendung in der empirischen Sozialforschung interessant zu machen. Zur Analyse dieser Modelle stehen mittlerweile eine Reihe von Computerprogrammen wie MLLSA (vgl. Eliason 1988; Clogg 1977) oder LCAG (Hagenaars 1987) zur Verfügung.

Bei der statistischen Analyse dieser Modelle geht es um die Schätzung der Wahrscheinlichkeiten für die Realisationen der latenten Klassen, der bedingten Wahrscheinlichkeiten unter der Bedingung einer latenten Klasse, eine bestimmte beobachtete Antwort gegeben zu haben, sowie der latenten Übergangswahrscheinlichkeiten. Bye/Schechter (1981) sowie Van de Pol/deLeeuw (1986) zeigen, wie man im Rahmen der LCA die Parameter latenter Markovketten schätzen kann. Van de Pol und Langeheine zeigen die Anwendung auf die Analyse stochastischer Prozesse, die sich aus heterogenen Teilprozessen zusammensetzen, wie z.B. Mover-Stayer-Modellen.

Diese Art der Analyse kann als kategoriales Analogon zur klassischen Analyse von Strukturgleichungsmodellen mit kontinuierlich verteilten latenten Variablen betrachtet werden. Anwendungen der LCA in der Panelanalyse sind bisher noch vergleichsweise selten.

3.2.2 Abschwächung von Verteilungsvoraussetzungen

Die zweite Form der Abschwächung von Modellannahmen betrifft die Abschwächung von Verteilungsvoraussetzungen bei der Schätzung von Parametern in Strukturgleichungsmodellen. Da diese Modelle einerseits von wesentlicher Bedeutung für die Panelmodellierung gerade in den Sozialwissenschaften sind, andererseits empirische Umfragedaten aber zumeist nicht normalverteilt sind, kommt dieser Form der Abschwächung besondere Bedeutung zu.

Inzwischen wurden zur Parameterschätzung Verfahren entwickelt, die Schätzungen unter verallgemeinerten Verteilungsbedingungen erlauben. Diese sogenannten ADF-Schätzer (asymptotically distribution-free, vgl. Bentler 1983; Browne 1982, 1984; Mooijart 1985; Shapiro 1983; Shapiro/Browne 1987) gestatten die Parameterschätzung unter der Bedingung, daß die Beobachtungsdaten elliptisch oder sogar arbiträr verteilt sind. So liefern diese Verfahren konsistente und effiziente Schätzungen sowie Standardfehler der Parameter auch dann, wenn die Daten z.B. schief verteilt sind oder wenn die Kurtosis der empirischen Verteilung von der Kurtosis der multivariaten Normalverteilung signifikant abweicht. Einige Programme wie EQS (Bentler 1985), LISREL VII (Jöreskog/Sörbom 1988) und LISCOMP (Muthén 1987) bieten inzwischen diese Schätzer an. Muthén/Kaplan (1985) zeigen, wie man diese Schätzer auf nicht-normalverteilte Likert-Variablen anwenden kann. Diese Verfahren basieren alle auf einem verallgemeinerten Gewichtetsten-Kleinste-Quadrate-Schätzer (weighted-least-squares, WLS), bei dem die Anpassungsfunktion

$$F = (\underline{s} - \underline{\sigma})' \underline{W}^{-1} (\underline{s} - \underline{\sigma})$$

minimiert wird. In dieser Gleichung ist \underline{s} der Vektor der beobachteten Stichprobenvarianzen, $\underline{\sigma}$ der Vektor der entsprechenden theoretischen Varianzen und Kovarianzen und \underline{W} ein konsistenter Schätzer der asymptotischen Kovarianzmatrix von \underline{s} . Über die Gewichtsmatrix lassen sich die Schätzungen unter verschiedenen Verteilungsbedingungen steuern (vgl. Faulbaum 1984, 1987). Diese neueren Schätzverfahren sind auch von großer Bedeutung für Populations- und Subgruppenvergleiche von Panelmodellen sowie für die Kohortenanalyse, wo nun auch unterschiedlichen Verteilungsvoraussetzungen in den verschiedenen Gruppen Rechnung getragen werden kann (vgl. Bentler/Lee/Weng 1987).

3.2.3 Verallgemeinerung von Parameterrestriktionen

Die dritte Abschwächung betrifft die Verallgemeinerung von Parameterrestriktionen. Ließen sich in LISREL lediglich Gleichheitsrestriktionen und nur mit Hilfe gewisser Kunstgriffe (vgl. Rindskopf 1983, 1984) auch allgemeinere Restriktionen testen, so erlauben einige Programme wie EQS, COSAN (Fraser ca. 1980) und LINCOS (Schoenberg/Arminger/Edlefsen 1987) bereits die Überprüfung allgemeinerer linearer und z.T. auch nicht-linearer Restriktionen. Dies hat besondere Konsequenzen für die Modellierung von Paneldaten, wo

sich nun z.B. allgemeine Hypothesen über die lineare oder nicht-lineare Abnahme und Zunahme von Stabilitäten und Reliabilitäten leichter testen lassen.

Alle bisher genannten Formen der Abschwächung betreffen auch die Analyse von latenten Mittelwertsstrukturen, wie sie etwa bei der Analyse von Wachstumskurven (vgl. z.B. McArdle 1986; McArdle/Epstein 1987) oder bei der Anwendung von Kohortenmodellen mit latenten Variablen in der Panelanalyse gefordert wird (vgl. Jöreskog/Sörbom 1985; Faulbaum 1987).

4. Abschließende Bemerkungen

Wie in anderen Bereichen der empirischen Sozialwissenschaften, so fällt auch im Bereich der Panelanalyse das Mißverhältnis zwischen der Vielfalt der theoretischen Weiterentwicklungen und der geringen Zahl ihrer praktischen Anwendungen auf. Viele empirische Studien begnügen sich mit speziellen, auf ihre eigene Fragestellung abzielenden, relativ einfachen statistischen Analysen. ZUMA hat in der Vergangenheit wiederholt versucht, durch spezielle Arbeitstagungen theoretische Weiterentwicklungen im Bereich der Panelforschung in die Praxis zu vermitteln und wird dies auch weiterhin tun. Auch im Bereich der ZUMA-Grundlagenforschung wird dieser Aspekt besonders akzentuiert werden. Von besonderer Bedeutung sind in diesem Zusammenhang die Konsequenzen der verschiedenen Formen der Abschwächung zugrundeliegender Modellannahmen für die Anwendbarkeit von Panelmodellen.

Für bestimmte Modellklassen stellt sich das Problem der Anwendbarkeit in verschärfter Form. Hierzu gehören insbesondere die zeitkontinuierlichen Modelle mit ihren speziellen Schätz-, Identifikations- und Einbettbarkeitsproblemen (vgl. Hamerle/Nagl/Singer 1988; Singer/Spilerman 1976b). Auch der von Tuma/Hannan (1984) ans Herz gelegte Übergang von der panelanalytischen zur ereignisanalytischen Betrachtungsweise scheint leider nur für bestimmte Gegenstandsbereiche plausibel oder möglich. Man kann zwar retrospektive Daten über Zeitpunkt und Dauer von Variablen, wie Arbeitslosigkeit, in Panelstudien miterheben, muß aber in der Regel davon ausgehen, daß diese Daten für Einstellungsvariablen nur sehr schwer zu erheben sind.

Im Einstellungsbereich empfiehlt es sich, möglichst präzise Hypothesen über die Kontextbedingungen für die Entstehung und Veränderung von Einstellungen zu formulieren. Zu diesen Kontextbedingungen gehören einerseits die zwischen den Beobachtungen möglicherweise aufgetretenen Einflüsse, die eine Wirkung auf die spätere Einstellung haben könnten, andererseits die Merkmale des Beobachtungszeitpunktes selbst. Referenzpunkt muß hier nicht eine unterstellte absolute Zeitachse sein, sondern ein Bündel von Annahmen über die Merkmalsstruktur der "Gelegenheiten", zu denen beobachtet wird. Besondere Aufmerksamkeit muß dabei der Konfundierung von Bedingungen (Effekten) gewidmet werden. Zur Separierung von Effekten bedarf es u.U. besonderer ex-

perimenteller oder quasi-experimenteller Designs. Es nimmt Wunder, daß in der einführenden sozialwissenschaftlichen Panelliteratur kaum auf den Designaspekt von Panelstudien eingegangen wird.

Ähnlich wie bei anderen Untersuchungsformen, ist auch hier die Generalisierbarkeit der Ergebnisse ein wichtiger Aspekt. Grundsätzlich muß davon ausgegangen werden, daß die im Rahmen von Panelstudien ermittelten Stabilitäten und Reliabilitäten nicht so ohne weiteres in die Zukunft generalisiert werden dürfen. Vermutungen darüber, ob ein gesellschaftlicher Prozeß einen Fixpunkt erreicht hat bzw. gewisse stationäre Eigenschaften aufweist, müssen reine Vermutungen bleiben, und dies aufgrund von Argumentationen, wie sie bei der Diskussion des Humeschen Kausalitätsbegriffs (Kausalität als "constant conjunction"; vgl. Bhaskar 1978; Baumrind 1983; Cummins 1983; Salmon 1984) vorgebracht wurden und die hier aus Platzgründen nicht aufgegriffen werden können. Mittlerweile dürfte aber vielen Sozialwissenschaftlern klar sein, daß sich kausale Beziehungen auch nicht im Rahmen von Panelstudien nachweisen lassen, auch dann nicht, wenn man über Spezifikations-tests und andere Verfahren statistisch nachgewiesen hat, daß es keine vermittelnden unbeobachteten Drittvariablen gibt. Dies bedeutet insbesondere, daß die Kausalität als Aspekt der Generalisierbarkeit von Panelergebnissen auf der Basis statistischer Datenanalyse keine Rolle spielen kann.

Für die Frage der Generalisierbarkeit scheint die Frage entscheidender zu sein, ob es einen theoretischen Rahmen gibt, dessen Reichweite über die Panelzeitpunkte hinausreicht und in den die in einer konkreten Studie gewonnenen Ergebnisse eingeordnet werden können. Gerade an der Spezifikation des Verhältnisses zwischen dem theoretischen Rahmen und der empirischen Realität mangelt es den meisten Studien. Dabei kann das eigentliche Ziel nicht darin liegen, diesen Rahmen auf der Basis der empirischen Daten eindeutig zu identifizieren, sondern eher darum zu prüfen, ob Klassen theoretischer Spezifikationen aufgrund der empirischen Daten ausgeschlossen werden müssen. Von daher muß die Bedeutung der Ergebnisse von Singer/Spilerman (1976a) relativiert werden.

Es liegt nahe, als theoretischen Rahmen ein zeitdiskretes oder zeitkontinuierliches Modell zu spezifizieren (vgl. hierzu auch Hannemann 1988). Dies reicht jedoch als theoretische Erklärung nicht aus, da auch die Art der Dynamik selbst erklärt werden muß. In diesem Zusammenhang sind Versuche zu sehen, die formalen Möglichkeiten theoretischer Spezifikationen zu erweitern und für die Datenanalyse nutzbar zu machen (vgl. Faulbaum 1986b). Ansatzpunkt ist hier die Erkenntnis, daß die logische Struktur sozialwissenschaftlicher Modelle und auch des Kausalitätsbegriffs eine breitere mathematische Struktur erfordern als die durch ein Gleichungssystem spezifizierbare (vgl. hierzu Glymour et al. 1987).

In vielen Fällen wird es sinnvoll sein, die Heterogenität der Population ins Auge zu fassen, separate Subgruppen zu analysieren und Vergleiche von Modellen über Subgruppen hinweg durchzuführen. Hier kommen insbesondere die erweiterten Möglichkeiten des Subgruppenvergleichs von Strukturgleichungsmodellen unter verallgemeinerten Verteilungsbedingungen und verallgemeinerten Parameterrestriktionen in Frage. Auch Subgruppenanalysen für diskrete Daten unter allgemeineren Parameterrestriktionen sind auf Grund von Fortschritten der LCA inzwischen prinzipiell möglich geworden (vgl. Clogg/Goodman 1985). Mitunter mag es auch hilfreich sein, in einer heuristischen Phase exploratorische Analysen zur Identifikation von Subgruppen vorzuschalten.

Dieser Beitrag wurde von Frank Faulbaum verfaßt.

Literatur

- Aigner, D.J./Hsiao, C./Kapteyn, A./Wansbeek 1983: Latent variables models in econometrics. S. 1321-1393 in Griliches, Z./Intriligator, M.D. (Eds.), Handbook of Econometrics. Vol. II. Amsterdam: North Holland.
- Amemiya, T. 1981: Qualitative response models: A survey. *Journal of Economic Literature* 19: 1483-1536.
- Amemiya, T. 1984: Tobit models: A survey. *Journal of Econometrics* 24: 3-61.
- Amemiya, T. 1985: *Advanced Econometrics*. Cambridge, Mass.: Harvard University Press.
- Anderson, T.W. 1979: Panels and Time Series Analysis: Markov Chains and Autoregressive Processes. S. 82-97 in Merton, R.K./Coleman, J.S./Rossi, P.H. (Eds.), *Qualitative and Quantitative Social Research*. New York: The Free Press.
- Anderson, T.W./Goodman, L.A. 1957: Statistical inference about Markov chains. *Annals of Mathematical Statistics* 28: 89-110.
- Anderson, T.W./Hsiao, C. 1982: Formulation and estimation of dynamic models using panel data. *Journal of Econometrics* 18: 47-82.
- Arminger, G. 1984: Modelltheoretische und methodische Probleme bei der Analyse von Paneldaten mit qualitativen Variablen. *Vierteljahrshefte zur Wirtschaftsforschung*, Heft 4: 470-480.
- Arminger, G. 1986: Linear stochastic differential equation models for panel data with unobserved variables. S. 187-212 in Tuma, N.B. (Hrsg.), *Sociological Methodology 1986*. San Francisco: Jossey-Bass.
- Arminger, G. 1987: Misspecification, asymptotic stability, and ordinal variables in the analysis of panel data. *Sociological Methods & Research* 15: 336-348.
- Arminger, G./Müller, F. in Vorbereitung: *Lineare Modelle zur Analyse von Paneldaten*.
- Arminger, G./Schoenberg, R.J. 1987: Pseudo maximum likelihood estimation and a test for misspecification in mean and covariance structure models. Eingereicht bei *Psychometrika*.
- Baumrind, D. 1983: Special causal attributions in the social sciences: the reformulated stepping-stone theory of heroin use as exemplar. *Journal of Personality and Social Psychology* 45: 1289-1298.
- Becker, R.A./Cleveland, W.S./Wilks, A.R. 1987: Dynamic graphics for data analysis. *Statistical Science* 2: 355-383.
- Bentler, P.M. 1980: Multivariate analysis with latent variables. *Annual Review of Psychology* 31: 419-456.
- Bentler, P.M. 1983: Some contributions to efficient statistics in structural models: specification and estimation of moment structures. *Psychometrika* 48: 493-517.
- Bentler, P.M. 1985: *Theory and Implementation of EQS. A Structural Equations Program*. BMDP Statistical Software, Inc.: Los Angeles, California.
- Bentler, P.M./Løe, S.Y. 1983: Covariance structures under polynomial constraints: Applications to correlation and alpha-type structural models. *Journal of Educational Statistics* 8: 207-222.
- Bentler, P.M./Lee, S.-Y./Weng, L.-J. 1987: Multiple population covariance structure analysis under arbitrary distribution theory. *Communications in Statistics. Theory and Methods*. Vol. 16: 1951-1964.
- Bentler, P.M./Weeks, D.G. 1980: Linear structural equations with latent variables. *Psychometrika* 45: 289-308.
- Bhargava, A./Sargan, J.D. 1983: Estimating dynamic random effects models from panel data covering short time periods. S. 9-44 in *Panel Data on Incomes*. London School of Economics.

- Bhaskar, R. 1978: A Realist Theory of Science. Hassocks/New Jersey: Harvester Press and Humanities Press.
- Bishop, Y.M.M./Fienberg, S.E./Holland, P.W. 1975: Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass.: the MIT Press.
- Blossfeld, H.-P./Hamerle, A./Mayer, K.U. 1986: Ereignisanalyse. Frankfurt: Campus.
- Blumen, I./Kogan, M./McCarthy, P.J. 1955: the Industrial Mobility of Labor as a Probability Process. Ithaca, New York: Cornell University Press.
- Blumen, I./Kogan, M./McCarthy, P.J. 1966: Probability models for mobility. S. 318-334 in Lazarsfeld, P.F./Henry, N.W. (Eds.), Readings in Mathematical Social Science. Cambridge, Mass.: The MIT Press.
- Bohrnstedt, G.W./Mohler, P.Ph./Müller, W. (Hrsg.) 1987: An Empirical Study of the Reliability and Stability of Survey Research Items. Sociological Methods & Research 15.
- Browne, M.W. 1982: Covariance structures. S. 72-141 in Hawkins, D.M. (Hrsg.), Topics in Appl Multivariate Analysis. London: Cambridge University Press.
- Browne, M.W. 1984: Asymptotically distribution-free methods for the analysis of covariance structures. British Journal of Mathematical and Statistical Psychology 37: 62-83.
- Bye, B.V./Schechter, E.S. 1986: A latent Markov model approach to the estimation of response errors in multivariate panel data. Journal of the American Statistical Association 81: 375-380.
- Carroll, G.R. 1982: Dynamic Analysis of Discrete Dependent Variables: A Didactic Essay. ZUMA-Arbeitsbericht Nr. 82/08.
- Chamberlain, G. 1982: Multivariate regression models for panel data. Journal of Econometrics 18: 6-46.
- Chamberlain, G. 1984: Panel data. S. 1243-1318 in Griliches, Z./Intriligator, (Eds.) Handbook of Econometrics, Vol. III. Elsevier Science Publishers.
- Chung, K.L. 1960: Markov Chains. New York: Springer Verlag.
- Clogg, C.C. 1977: Unrestricted and Restricted Maximum Likelihood Latent Structure Analysis: A Manual for Users. Working Paper No. 1977-09, Population Issues Research Center, The Pennsylvania State University.
- Clogg, C.C. 1981: Latent structure models of mobility. American Journal of sociology 86: 836-868.
- Clogg, C.C./Goodman, L.A. 1985: Simultaneous latent structure analysis in several groups. In Tuma, N.B. (Ed.), Sociological Methodology 1985. San Francisco: Jossey Bass.
- Coleman, J.S. 1964: Models of Change and Response Uncertainty. Englewood Cliffs: Prentice Hall.
- Coleman, J.S. 1968: The mathematical study of change. S. 428-478 in Bialock, H.M./Bialock, A.B. (Eds.), Methodology in Social Research. New York: McGraw-Hill.
- Coleman, J.S. 1981: Longitudinal Data Analysis. New York: Basic Books, Inc., Publ.
- Collins, L.M./Cliff, N. 1985: Axiomatic foundations of a three-set Guttman simplex model with applicability to longitudinal data. Psychometrika 50: 147-158.
- Cummins, R. 1983: The Nature of Psychological Explanation. Cambridge, Mass.: MIT Press.
- Dayton, C.M./Macready, G.B. 1983: Latent structure analysis of repeated classifications with dichotomous data. British Journal of Mathematical and Statistical Psychology 36: 189-201.
- Dhrymes, P.J. 1984: Limited Dependent Variables. S. 1567-1631 in Griliches, Z./Intriligator, M.D. (Hrsg.), Handbook of Econometrics. Vol. III. Elsevier Science Publishers.
- Diekmann, A. 1988: Ereignisdatenanalyse - Beispiele, Probleme und Perspektiven (in diesem Heft).
- Diekmann, A./Mitter, P. 1984: Methoden zur Analyse von Zeitverläufen. Stuttgart: Teubner.
- Duncan, O.D. 1975: Introduction to Structural Equation Models. New York: Academic Press.
- Duncan, G.J./Morgan, J.N. 1984: Behavioral research with the panel study of income dynamics in retrospect and prospect. Vierteljahrshefte zur Wirtschaftsforschung, Heft 4: 415-427.
- Edlefsen, L.E./Jones, S.D. 1986: GAUSS Programming Manual. Kent, Wa.: Aptech Systems.
- Ellison, S.R. 1988: The Categorical Data Analysis System. Version 3.00. User's Manual. Department of Sociology and Population Issues Research Center. The Pennsylvania State University.
- Faulbaum, F. 1984: EQS: BMDP's Antwort auf LISREL. ZUMA-Nachrichten 15: 72-82.
- Faulbaum, F. 1986a: Programme zur Anpassung von Strukturgleichungsmodellen mit latenten Variablen: Eine vergleichende Darstellung von EQS und LISREL. S. 165-177 in Lehmacher, W./Hörmann, A. (Hrsg.), Statistik-Software. Stuttgart/New York: Gustav Fischer.
- Faulbaum, F. 1986b: Very Soft Modeling. ZUMA-Arbeitsbericht Nr. 86/04.
- Faulbaum, F. 1987: Intergroup comparisons of latent means across waves. Sociological Methods & Research 15: 317-335.
- Formann, A.K. 1984: Die Latent-Class-Analyse. Weinheim/Basel: Beltz.
- Fraser, C. undated: COSAN User's Guide. Ontario: The Ontario Institute for Studies in Education.
- Frederiksen, C.H./Rotondo, J.A. 1979: Time-series models and the study of longitudinal change. S. 111-153 in Nesselroade, J.R./Baltes, P.B. (Eds.), Longitudinal Research in the Study of Behavior and Development. New York: Academic Press.

- Glymour, C./Scheines, R./Spirtes, P./Kelly, K. 1987: *Discovering Causal Structure*. New York: Academic Press.
- Hagenaars, J.A. 1987: LCAG: loglinear modelling with latent variables: A Modified LISREL approach. In: Saris, W./Gallhofer, I. (Hrsg.), *Sociometric Research*. Vol. 2. London: MacMillan.
- Hagenaars, J.A. 1988: Latent structure models with direct effects between indicators. *Sociological Methods & Research* 16: 379-405.
- Hagenaars, J.A./Luikx 1987: LCAG: Latent Class Models and Other Loglinear Models with Latent Variables. User's Manual. Working Paper Series, Tilburg University, Department of Sociology.
- Hajek, P./Havranek, T. 1978a: *Mechanizing Hypothesis Formation*. Berlin/Heidelberg/New York: Springer.
- Hajek, P./Havranek, T. 1978b: The GUHA-method - its aims and techniques. *International Journal of Man-Studies* 10: 3-22.
- Hamerle, A./Nagl, W./Singer, H. 1988: Parameter estimation in continuous time dynamic systems using panel data. Diskussionsbeiträge Nr. 109/s. Fakultät für Wirtschaftswissenschaften und Statistik der Universität Konstanz.
- Hanefeld, U. 1984: Das Sozio-ökonomische Panel. Eine Längsschnittstudie für die Bundesrepublik Deutschland. Vierteljahrshefte zur Wirtschaftsforschung, Heft 4: 391-406.
- Hanefeld, U. 1987: Das Sozio-ökonomische Panel. Grundlagen und Konzeption. Frankfurt/Main/New York: Campus.
- Hanneman, R.A. 1988: *Computer-Assisted Theory Building*. Beverly Hills: Sage.
- Hausman, J.A. 1978: Specification tests in econometrics. *Econometrica* 46: 1251-1272.
- Hausman, J.A. 1983: Specification and estimation of simultaneous equation models. S. 391-448 in Griliches, Z./Intriligator, M.D. (Eds.), *Handbook of Econometrics*. Vol. II. Amsterdam: North-Holland.
- Hausman, J.A./Taylor, W.E. 1981: Panel data and unobservable individual effects. *Econometrica* 49: 1377-1398.
- Heckman, J. 1978a: Dummy endogeneous variables in a simultaneous equation system. *Econometrica* 46: 931-959.
- Heckman, J. 1978b: Simple statistical models for discrete panel data developed and applied to test the hypothesis of true state dependence against the hypothesis of spurious state dependence. *Annales de l'INSEE* 30-31: 227-269.
- Heckman, J. 1981: Statistical models for discrete panel data. S. 179-195 in Manski, C.F./McFadden, D. (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, Mass.: MIT Press.
- Heckman, J./Singer, B. 1984: Econometric duration analysis. *Journal of Econometrics* 24: 63-132.
- Heckman, J./Singer, B. 1986: Econometric analysis of longitudinal data. S. 1689-1763 in Griliches, Z./Intriligator, M.D. (Hrsg.), *Handbook of Econometrics*, Vol. III. Elsevier Science Publishers.
- Heise, D.R. 1969: Separating reliability and stability in test-retest-correlation. *American Sociological Review* 34: 93-101.
- Hsiao, C. 1986: *Analysis of Panel Data*. Cambridge: Cambridge University Press.
- Huber, P.J. 1985: Projection pursuit. *Annals of Statistics* 13: 435-475.
- Jöreskog, K.G. 1969: A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34: 183-202.
- Jöreskog, K.G. 1970: Estimation and testing of simplex models. *The British Journal of Mathematical and Statistical Psychology* 23: 121-145.
- Jöreskog, K.G. 1973: A general method for estimating a linear structural equation system. S. 85-112 in Goldberger, A.S./Duncan, O.D. (Eds.), *Structural Equation Models in the Social Sciences*. New York: Seminar Press.
- Jöreskog, K.G. 1974: Analyzing Psychological Data by Structural Analysis of Covariance Matrices. S. 1-56 in Krantz, D.H./Atkinson, R.C./Luce, R.D./Suppes, P. (Eds.) *Measurement, Psychophysics, and Neural Information Processing*. Contemporary Developments in Mathematical Psychology. Vol. II. San Francisco: W.H. Freeman.
- Jöreskog, K.G. 1978: Structural analysis of covariance and correlation matrices. *Psychometrika* 43: 443-477.
- Jöreskog, K.G. 1979: Statistical estimation of structural models in longitudinal-developmental investigations. S. 303-351 in Hesselrode, J.R./Baltes, P.B. (Hrsg.) *Longitudinal Research in the Study of Behavior and Development*. New York: Academic Press.
- Jöreskog, K.G. 1981: Statistical models for longitudinal studies. S. 118-124 in Schulsinger, F./Mednick, S.A./Knop, J. (Hrsg.) *Longitudinal research*. Boston/The Hague/London: Martinus Nijhoff Publishing.
- Jöreskog, K.G./Sörbom, D. 1985: Simultaneous analysis of longitudinal data from several cohorts. S. 232-341 in Mason, W.M./Fienberg, S.E. (Hrsg.), *Cohort Analysis in Social Research*. New York/Berlin/Heidelberg: Springer.
- Jöreskog, K.G./Sörbom, D. 1988: *LISREL VII*. New York: John Wiley.

- Kessler, R.C./Greenberg, D.F. 1981: Linear Panel Analysis. New York: Academic Press.
- Klevmarken, N.A. 1984: Household market and nonmarket activities. Vierteljahrshefte für Wirtschaftsforschung, Heft 4: 452-457.
- Krupp, H.-J./Hanefeld, U. (Hrsg.) 1987: Lebenslagen im Wandel: Analysen 1987. Frankfurt/M./New York: Campus.
- Langeheine R. 1987: Black & White, anfängliche Antwortunsicherheit, Mover-Stayer, Third Force oder was? Ein paar weitere Überlegungen zu Jagodzinski's Analyse des Postmaterialismus-Panels. ZA-Information 20: 44-55.
- Lazarsfeld, P.F. 1950: The logical and mathematical foundation of latent structure analysis. Chapter 10 in Stouffer, S.A. et al. (Hrsg.), Measurement and Prediction. Princeton: Princeton University Press.
- Lazarsfeld, P.F. 1960: Latent Structure Analysis. S. 476-535 in Koch, S. (Hrsg.), Psychology: A Study of Science. New York: McGraw-Hill.
- Lazarsfeld, P.F. 1965: Latent structure analysis. S. 37-54 in Sternberg, S. et al. (Hrsg.), Mathematics and Social Sciences. The Hague: Mouton.
- Lazarsfeld, P.F./Henry, N.W. 1968: Latent Structure Analysis. Boston: Houghton Mifflin Company.
- Little, R.J.A./Rubin, D.B. 1987: Statistical Analysis with Missing Data. New York: Wiley.
- Maddala, G.S. 1983: Limited Dependent Variables in Econometrics. Cambridge: Cambridge University Press.
- Marini, M.M./Olson, A.R./Rubin, D.B. 1980: Maximum likelihood estimation in panel studies with missing data. S. 314-357 in Schuessler, K. (Ed.), Sociological Methodology 1980. San Francisco: Jossey Bass.
- Markus, G.B. 1979: Analyzing Panel Data. Beverly Hills: Sage.
- McCutcheon, A.L. 1987: Latent Class Analysis. Beverly Hills: Sage.
- McDonald, R.P. 1980: A simple comprehensive model for the analysis of covariance structures: Some remarks and applications. British Journal of Mathematical and Statistical Psychology 33: 161-183.
- McFadden, D.L. 1984: Econometric analysis of qualitative response models. S. 1396-1457 in Griliches, Z./Intriligator, M.D. (Hrsg.), Handbook of Econometrics. Vol. III. Elsevier Science Publishers.
- McArdle, J.J./McDonald, R.P. 1984: Some algebraic properties of the reticular action model for moment structures. British Journal of Mathematical and Statistical Psychology 37: 234-251.
- McArdle, J.J. 1986: Latent growth within behavior genetic models. Behavior Genetics 16: 163-200.
- McArdle, J.J./Epstein, D. 1987: Latent growth curves within developmental structural equation models. S. 110-133 in Connell, J.P./Tanaka, J.S. (Hrsg.), Special Section on Structural Equation Modeling. Child Development 58, Heft 1.
- Möbus, C. 1986: Analyse von Rohdaten mit LISREL - Allgemeines Lineares Modell, stochastische Differenzen- und Differentialgleichungssysteme.
- Möbus, C./Schneider, W. 1986: Strukturmodelle für Längsschnittdaten und Zeitreihen. Bern/Stuttgart/Toronto: Verlag Hans Huber.
- Möbus, C./Nagl, W. 1983: Messung, Analyse und Prognose von Veränderungen. S. 239-470 in Bredenkamp, J./Feger, H. (Hrsg.) Hypothesenprüfung. Göttingen: Verlag für Psychologie Dr. C.J. Hogrefe.
- Mooijlaart, A. 1985: Factor analysis for non-normal variables. Psychometrika 50: 323-342.
- Muthén, B.O. 1978: Contributions to factor analysis of dichotomous variables. Psychometrika 43: 551-560.
- Muthén, B.O. 1982: Some categorical response models with continuous latent variables. S. 65-79 in Jöreskog, K.G./Wold, H. (Hrsg.), Systems under Indirect Observation: Causality, Structure, Prediction. Amsterdam: North Holland.
- Muthén, B.O. 1983: Latent variable structural modeling with categorical data. Journal of Econometrics 22: 43-65.
- Muthén, B.O. 1984: A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika 49: 115-132.
- Muthén, B.O. 1987: LISCOMP. Analysis of Linear Structural Relations Using a Comprehensive Measurement Model. Scientific Software, Mooresville, USA.
- Muthén, B.O./Christofferson, A. 1981: Simultaneous factor analysis of dichotomous variables in several groups. Psychometrika 46: 485-500.
- Muthén, B.O./Kaplan, D. 1985: A comparison of some methodologies for the factor analysis of non-normal Likert variables. British Journal of Mathematical and Statistical Psychology 38: 171-189.
- Olsson, U. 1979: Maximum likelihood estimation of the polychoric correlation coefficient. Psychometrika 44: 443-460.
- Olsson, U./Drasgow, F./Dorans, N.J. 1982: The polyserial correlation coefficient. Psychometrika 47: 337-347.
- Plenis, I. 1985: Analyzing Change. New York: Wiley.

- Raffalovich, L.E./Bohrnstadt, G.W. 1987: Common, specific and error variance components of factor models: estimation with longitudinal data. *Sociological Methods & Research* 15: 385-405.
- Rindskopf, D. 1983: Parameterizing inequality constraints on unique variances in linear structural models. *Psychometrika* 48: 73-83.
- Rindskopf, D. 1984: Using phantom variables and imaginary variables to parameterize constraints in linear structural models. *Psychometrika* 49: 37-47.
- Rubin, D.B. 1987: *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- Salmon, H.C. 1984: *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Sawitzky, G. 1988: Graphische Datenanalyse mit MacSpin. S. 361-369 in Faulbaum, F./Uehlinger, H.-M. (Hrsg.), *Fortschritte der Statistik-Software I*. Stuttgart: Gustav Fischer.
- Schaber, G./Born, M./Dickes, P./Gailly, B./Hausman, P. 1982: Processus de pauperisation dans les groupes a risque eleve de deprivation. Approfondissement de l'etude transnationale sur la pauvrete persistante. CEPS. Luxembourg.
- Singer, B. 1982: Aspects of nonstationarity. *Journal of Econometrics* 18: 169-190.
- Singer, B./Spilerman, S. 1976a: The representation of social processes by Markov models. *American Journal of Sociology* 82: 1-54.
- Singer, B./Spilerman, S. 1976b: Some methodological issues in the analysis of longitudinal surveys. *Annals of Economic and Social Measurement* 5: 447-472.
- Shapiro, A. 1983: Asymptotic distribution theory in the analysis of covariance structures. *South African Statistical Journal* 17: 33-81.
- Shapiro, A./Browne, M.W. 1987: Analysis of covariance structures under elliptical distributions. *Journal of the American Statistical Association* 82: 1092-1097.
- Schoenberg, R./Arminger, G./Edlefsen, L. 1987: *LINCS. Linear Covariance Structure Analysis*. RJS Software, Kensington, Maryland.
- Soong, T.T. 1973: *Random Differential Equations in Science and Engineering*. New York: Academic Press.
- Tobin, J. 1958: Estimation of relationships for limited dependent variables. *Econometrica* 26: 24-36.
- Tuma, N.B./Hannan, M.T. 1984: *Social Dynamics*. New York: Academic Press.
- Van de Pol, F./De Leeuw, J. 1986: A latent Markov model to correct for measurement error. *Sociological Methods & Research* 15: 118-141.
- Van de Pol, F./Langeheine, R. 1988: Mover-Stayer Models, Mixed Markov Models and the EM Algorithm. Central Bureau of Census, Department of Statistical Methods, Voorburg, The Netherlands.
- Van de Stadt, H. 1987: The dynamics of income distribution. Paper presented at the workshop on life events and income change, CEPS, Walferdange.
- Van de Stadt, H./De Kleyen 1986: Income distribution trends in the Netherlands, 1959-1982. Netherlands Central Bureau of Statistics.
- Van de Stadt, H./Ten Cate, A./Hundepool, A.J./Keller, W.J. 1986: Koopkracht in kaart gebracht, een statistiek van de inkomensdynamiek (Real income analyzed, a statistical study of income dynamics). Statistische onderzoeken M28, Netherlands Central Bureau of Statistics.
- Van der Heijden, P.G.M. 1987: *Correspondence Analysis of Longitudinal Categorical Data*. Leiden: DSWO Press.
- White, H. 1985: Tests for misspecification. S. 552-555 in Kotz, S./Johnson, N.L. (Eds.), *Encyclopedia of Statistical Science Vol. 5*. New York: Wiley.
- Wiggins, L.M. 1951: The application of latent structure analysis to the problem of reliability. Part III of Lazarsfeld, P.F. (Hrsg.), *The Use of Mathematical Models in the Measurement of Attitudes*. Santa Monica: the Rand Corporation.
- Wiggins, L.M. 1973: *Panel Analysis*. Amsterdam: Elsevier Scientific Publishing Company.
- Wiley, D.E./Wiley, J.A. 1970: The estimation of measurement error in panel data. *American Sociological Review* 35: 112-117.
- Young, F.W. 1988: VISUALS: Software for dynamic hyper-dimensional graphics. S. 331-348 in Faulbaum, F./Uehlinger, H.-M. (Hrsg.), *Fortschritte der Statistik-Software I*. Stuttgart: Gustav Fischer.